(19) JAPANESE PATENT OFFICE

PATENT ABSTRACTS OF JAPAN

(11) Publication number: **06019861 A**
(43) Date of publication of application: **28.01.1994**

(51) Int. Cl **G06F 15/16**

(21) Application number: **05087579**
(22) Date of filing: **14.04.1993**
(30) Priority: **30.04.1992 US 92 876670**

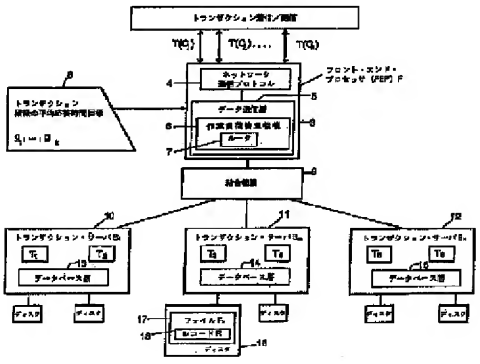
(71) Applicant: **INTERNATL BUSINESS MACH CORP <IBM>**
(72) Inventor: **FERGUSON DONALD F**
GEORGIADIS LEONIDAS
NIKOLAOU CHRISTOS N

(54) MECHANISM FOR DESIGNATION PATH OF TRANSACTION PROCESSING SERVER

(57) Abstract:

PURPOSE: To adjust the load on each transaction processing server to attain the best responding performance in relation with an actual processing time across all hierarchies by classifying a transaction into plural hierarchies based on the request processing time, and setting a target responding time for every hierarchy in a multiprocessor transaction processing system.

CONSTITUTION: A load adjusting mechanism calculates a mean responding time for every present hierarchy, and the performance index of each hierarchy is calculated by considering a relation with the set target responding time of the hierarchy. When a new transaction arrives, the load adjusting mechanism inspects several transaction servers being several candidates whose paths can be designated by the transaction, and the performance index of every hierarchy when each is turned into the server is predicted. Then, an overall target established index obtained by synthesizing the performance indexes of all the hierarchies is derived, and the property of the transaction processing server whose path should be designated is evaluated.



(19)日本国特許庁（ＪＰ）　(12) **公開特許公報（A）**　(11)特許出願公開番号

特開平6－19861

(43)公開日　平成６年(1994)１月28日

(51)Int.Cl.[5]	識別記号	庁内整理番号	ＦＩ	技術表示箇所
Ｇ０６Ｆ　15/16	３８０　Ｚ	8840－5L		

審査請求　有　　請求項の数４(全 16 頁)

(21)出願番号　特願平5－87579

(22)出願日　平成５年(1993)４月14日

(31)優先権主張番号　８７６６７０

(32)優先日　1992年４月30日

(33)優先権主張国　米国（ＵＳ）

(71)出願人　390009531
インターナショナル・ビジネス・マシーンズ・コーポレイション
ＩＮＴＥＲＮＡＴＩＯＮＡＬ　ＢＵＳＩＮＥＳＳ　ＭＡＳＣＨＩＮＥＳ　ＣＯＲＰＯＲＡＴＩＯＮ
アメリカ合衆国10504、ニューヨーク州　アーモンク　（番地なし）

(72)発明者　ドナルド・フランシス・ファーガソン
アメリカ合衆国11364、ニューヨーク州ベイサイド、トゥハンドレッド・ナインティーンス・ストリート　61－29

(74)代理人　弁理士　頓宮　孝一　（外４名）

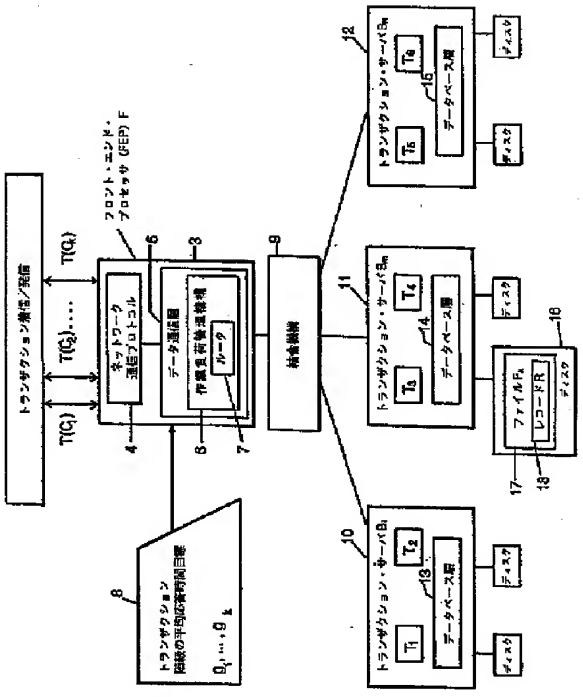
最終頁に続く

(54)【発明の名称】　トランザクション処理サーバの経路指定機構

(57)【要約】

【目的】　マルチプロセッサ・トランザクション処理システムにおいてトランザクションをその要求処理時間にもとづいて複数の階級に分類し、前記階級ごとに目標応答時間を設定し、階級全体について実際の処理時間との関係で最も応答性能がよくなるように各トランザクション処理サーバの負荷を調整する。

【構成】　負荷調整機構は現在の階級ごとの平均応答時間を算出し、その階級の設定目標応答時間との関係を考慮しつつ階級ごとの性能指標を算出する。新たなトランザクションが到着した時は、負荷調整機構はそのトランザクションが経路指定されることができるいくつかの候補となるトランザクション・サーバを検討し、各々がサーバとなった時の階級ごとの性能指標を予測する。次に全ての階級の性能指標を統合した全体目標達成指標を導出した上で経路指定すべきトランザクション処理サーバの適否を評価する。



1

【特許請求の範囲】

【請求項１】複数の階級に区別されたトランザクションを単位として処理を行う複数のトランザクション処理サーバを有しているコンピュータ・システムにおいて、最適な応答時間を確保するために前記トランザクションを前記トランザクション処理サーバに経路指定する機構であって、

前記各々の階級ごとに現在の平均応答時間を計算する手段と、

前記各々の階級ごとに目標応答時間を記憶する手段と、

前記現在の平均応答時間と前記目標応答時間とにもとづいて、前記各々の階級ごとの現在の性能指標を求める手段と、

１の着信トランザクションに応答し、前記着信トランザクションを処理するために経路指定されうる全ての前記トランザクション処理サーバごとに、前記着信トランザクションが前記指定されうる前記トランザクション処理サーバのうちの１つにおいて処理される場合の、前記各々の階級ごとの前記性能指標を予測して予測性能指標を求める手段と、

前記予測性能指標に対応して、それらを全階級について統合した全体目標達成指標を前記経路指定されうる全てのトランザクション処理サーバごとに予測する手段と、

前記全体目標達成指標のうち最良のものを決定する手段と、

前記決定に対応する前記トランザクション処理サーバに前記着信トランザクションを経路指定する手段と、

を含む経路指定機構。

【請求項２】１の階級に係わる前記階級ごとの現在の平均応答時間は、前記１の階級において処理が完了したトランザクションの到着時間と処理終了時間の差分値の時間重み付けされた平均値であることを特徴とした請求項１の経路指定機構。

【請求項３】前記予測性能指標を求める手段は、前記着信トランザクションの応答時間を予測する応答時間予測手段を有することを特徴とした請求項１の経路指定機構。

【請求項４】前記階級について前記現在の性能指標にもとづいて優先順位を付与し、その順に並べる手段と、

前記各々のトランザクション処理サーバに経路指定された前記トランザクションを前記優先順位の順に処理する手段と、

をさらに含む請求項１の経路指定機構。

【発明の詳細な説明】

【０００１】

【産業上の利用分野】本発明は、一般にトランザクションを処理する相互接続されたコンピュータ・システムにおける作業負荷管理プログラムに関し、特に各々のトランザクション階級が個々の応答時間目標を持つ、異なる階級のトランザクションを有するコンピュータ・システムにおける作業負荷管理プログラムに関する。

2

【０００２】

【従来の技術】トランザクション処理システムは、一般に広範囲の地理的区域に分散させられた複数のターミナルをサポートする多重プロセッサ・データ処理システムで実行する、オン・ラインのアプリケーション指向のシステムである。典型的なトランザクション処理システムは、ＩＢＭのＡＣＰ（エアライン制御プログラム）であり、航空会社予約システムで従来から使用されているが、また他のシステムでも使用され、特に銀行によってオン・ラインのテラー・アプリケーションにおいて使用されている。

【０００３】従来の多量の作業は多重プロセッサ及び分散処理システムにおいて、作業負荷管理プログラムによって実行された。静的及び動的なロード・バランス・アルゴリズムに関しては、著者D. L. Eager、et al. による題名"Adaptive LoadSharing in Homogeneous Distributed Systems"（IEEE Transactions onSoftware Engineering、SE-12 (5) 、May 1986）、著者Y. T. Wang、et al. による題名"Load Sharing in Distributed Systems"（IEEE Transactions onComputers、C-34 (3) 、March 1985）、及び著者A. N. Tantawi、et al. による題名"Optimal Static Load Balancing in Distributed Computer Systems、（Journal of the ACM、pages 445-465、April 1985）で述べられているので参照されたい。しかしながら、上記のアルゴリズムは着信する全てのトランザクションの平均応答時間を最小にするのにとどまり、各々が異なる応答時間目標を有するトランザクションの異なる階級を考慮していない。

【０００４】著者P. P. Bhattacharya、et al. による題名"Optimality and Finite TimeBehavior of an Adaptive Multi-objective Scheduling Algorithm"（Proc. ofthe 29th Conf. on Decision and Control、Honolulu、Hawaii、Dec. 1990）においては、ジョブ階級は割当てられた応答時間目標に該当し、スケジューリング・アルゴリズムはジョブをスケジュールするために提案されているが、これら全ては単一のプロセッサ内で実行される。上記記載のアルゴリズムは多重プロセッサ・システムへの応用においてどのように拡張されるのか或いは修正されるのか、及び複数のコンピュータに対して異なる応答時間階級目標を有するトランザクションの経路指定の追加された問題をどのように処理するかが示されていない。

【０００５】静的及び動的なトランザクションの経路指定アルゴリズムの研究が行われ、その結果が、著者P. S. Yu、et al. による題名"Dynamic Transaction Routing inDistributed Database Systems"（IEEE Transactions on SoftwareEngineering、SE-14 (9) 、September 1988）及び著者D. Cornell、et al. による題名"On Coupling Partitioned Database Systems"（Proceedings o

3

f 6thInternational Conference on Distributed Computing Systems、1986）で報告されているので参照されたい。しかしながら、上記におけるトランザクション経路指定の研究は、また全着信トランザクションにおける平均応答時間を最小にするのにとどまり、階級化された応答時間を前提としていない。一般に全体的な応答時間の最小化は、階級の特定の目標を満足させない。

【０００６】トランザクションの異なる階級のためには、異なる応答時間目標に従って作業負荷のバランスをとる、トランザクション多重処理システムの作業負荷管理プログラムを有することが望ましい。

【０００７】

【発明が解決しようとする課題】本発明の目的は、応答時間が目標に達成しない程度を最小にするための作業負荷のバランスをとる相互接続されたトランザクション処理コンピュータ・システムのための作業負荷管理プログラムを提供することにある。

【０００８】本発明の目的は、トランザクションが階級に分類され、各々の階級が個々の応答時間目標を持つシステムのための作業負荷管理プログラムを提供することにある。

【０００９】本発明の目的は、個々の応答時間目標に対して性能不十分であったトランザクションの階級における応答時間未達成性能指標を最小にするように、上記システムのプロセッサ間において、作業負荷のバランスがとれるような作業負荷管理プログラムを提供することにある。

【００１０】本発明の目的は、トランザクションの異なる階級の異なる応答時間目標を考慮しつつ、トランザクションの経路指定管理によって作業負荷のバランスをとることにある。

【００１１】本発明の目的は、複数の経路指定をとった場合の各着信トランザクションの応答時間をそれぞれ予測することにある。

【００１２】本発明の目的は、トランザクションのスケジューリングについて各階級毎に優先順位という概念を導入し、応答時間未達成性能指標に従ってこれらのスケジューリング優先順位を動的に調整することによって、作業負荷のバランスをとることにある。

【００１３】本発明の目的は、スケジューリング優先順位を動的に調整することに伴う障害を考慮するために、トランザクションが経路指定された目的のプロセッサにおけるいかなるマルチプログラミングをも適用可能な経路指定管理によって作業負荷のバランスをとることにある。

【００１４】本発明の目的は、特定の階級のトランザクションが必要とする記録に対して、より遅い平均アクセス時間をもつ他のプロセッサよりも、より速い平均アクセス時間を有するプロセッサによって、特定の階級のトランザクションがより能率的に実行できることを事前に

4

考慮して作業負荷のバランスをとることにある。

【００１５】

【課題を解決するための手段】上記目的及び利点は、トランザクションを処理する複数のサーバと着信トランザクションをサーバに分散させるための作業負荷管理プログラムを具備する本発明によって得られる。トランザクションをその応答時間目標によって、いくつかの階級に分類する。

【００１６】トランザクション階級はトランザクション・プログラム名、トランザクション・コード、ユーザ、トランザクションを提示するターミナル、またはワークステーション、及び多数の他の要素にもとづいて分類されたトランザクションのグループである。階級応答時間目標はオペレータによって外部から指定する等の方法で設定される。

【００１７】作業負荷管理プログラムは各トランザクション階級の現在の階級平均応答時間（average class response time ）を計算し、各階級目標応答時間との関連において、各階級の現在の階級性能指標を導き出す。好ましい実施例では、ある階級の現階級性能指標は（階級平均応答時間）（階級応答時間目標）によって表される。トランザクション階級の現平均階級応答時間及び各階級の現階級性能指標は、トランザクションが処理完了するたびに更新される。

【００１８】トランザクションが着信する度に作業負荷管理プログラムは、着信トランザクションがとりうる多数の経路を熟慮し、可能な経路のそれぞれについて階級性能指標を予測する。好ましい実施例では、全ての可能な経路が熟慮される。全資源消費量、全Ｉ／Ｏ遅延時間、及び全通信遅延時間がそれぞれ経路指定に対して予測される。全体目標達成指標が各々の経路に対して求められ、最良の全体目標達成指標が得られる経路の選択が行われる。好ましい実施例では、可能な経路指定選択の全体的な目標達成指標は該経路指定選択の予測された最悪（すなわち最高位）の階級性能指標であり、最良の全体的な目標達成指標は、これらの最悪の階級性能指標の最良（すなわち、最低位）の１つである。これは最小／最大機能と呼ばれることがある。

【００１９】好ましい実施例では、作業負荷管理プログラムは、また性能の良くない階級のトランザクションが、経路指定された宛先のサーバにおいて、性能の良い階級のトランザクションよりもさらに高位のタスク指名順位が得られるように、現階級性能指標、現平均階級応答時間と階級応答時間目標との比率に従って階級を優先設定する。

【００２０】各々のサーバは互いに機能的に独立しており、別々のコンピュータで実行される。所望するならば、物理的に同じコンピュータによって複数のサーバを実行できる。ルータ機能は、また各サーバから機能的に独立しており、別々のコンピュータまたは１つ以上のサ

5

ーバを実行するコンピュータによっても実行できる。

【００２１】

【実施例】

【数１】

$$\vec{N}$$

は以降ベクトルＮと記載する。本発明は、多重プロセッサ・トランザクション処理システムにおいて、高性能の動的な作業負荷管理を実行するために新しい技術を実施する。各トランザクションは一般にユーザ入力データを処理し、１つ以上のデータベースに対して読出し及び更新を実行し、アプリケーション・コードを実行してその結果をユーザに送信する。ユーザと実行中のトランザクションとの間で複数の対話ができる。

【００２２】システム・ハードウェア・アーキテクチャの好ましい実施例が、図１に示されている。システムで実行される作業は、データベース要求を実行するユーザ提示のトランザクションである。著者C. J. Date、による題名"An Introduction toDatabase Systems" (Volume II、Addison Wesley Company、Reading Mass. 、1983) は、典型的なトランザクション及びデータベース要求実行を説明しているので参照されたい。各々の提示されたトランザクションは予め定義されたトランザクション階級に属する。図１は６つの異なるトランザクションT_1、T_2、T_3、T_4、T_5、T_6を示し、トランザクション・サーバB_1、トランザクション・サーバB_n及びトランザクション・サーバB_Nによって実行される。C_1、C_2、．．．、C_Kで表されるＫ個のトランザクション階級がある。階級C_1の着信または発信のトランザクションは、例えばＴ（C_1）として表される。ある階級での着信トランザクションの識別は、トランザクション・プログラム名、トランザクション・コード、ユーザ及び多数の他の要素に依存する。各々のトランザクションはユーザ入力データを処理し、データベースに対して読出し及び更新を実行し、アプリケーション・コードを実行してその結果をユーザに送信する。ユーザと実行中のトランザクションとの間では複数の対話ができる。

【００２３】トランザクション処理システムはフロント・エンド・プロセッサ（ＦＥＰ）Ｆ及びＮ個のトランザクション・サーバB_1、B_2、B_n、．．．、B_Nを有する。各々のサーバは完全に機能的なコンピュータ・システムであり、ＣＰＵ、メモリ、ディスク、その他を有する。これらのコンピュータは結合機構９を介して、例えば、高帯域幅相互接続ネットワーク等に相互接続される。Ｆで表されるプロセッサはフロント・エンド・プロセッサ（ＦＥＰ）として指定され、ルータ７を含む作業負荷管理プログラム６を有する。ＦＥＰはユーザとの通信に必要なネットワーク通信プロトコル４を実行する。ＦＥＰは、またデータベース・システムの機能であるデータ通信層５を実行する。著者C. J. Date、による題

6

名"AnIntroduction to Database Systems" (Volume II、Addison Wesley Company、Reading Mass. 、1983) は、データ通信層がソフトウェアとどのようにインタフェースするかを述べている。このソフトウェアはデータベースのユーザによって使用されるターミナル、またはインテリジェント・ワークステーションとのメッセージのやりとりをするために送受信するプロセッサの通信ハードウェアを制御する。ユーザによって発信されたメッセージは、データ通信層によって特定の種類のトランザクションを開始させる。実行中のトランザクションから発信されるユーザへの応答がメッセージに組込まれてユーザに送信される。データ通信層５の構成要素の１つは作業負荷管理プログラム６である。作業負荷管理プログラム６の機能の１つは、例えばT_1、．．．、T_6を始めとするトランザクションを適切なトランザクション・サーバに経路指定することにある。これは、作業負荷管理プログラムの構成要素であるルータ７によって実行される。全ての着信及び処理完了のトランザクションはＦＥＰを通る。本明細で述べる技術は、一般に複数のＦＥＰの場合である。

【００２４】その他のプロセッサは、バック・エンド・プロセッサ（ＢＥＰ）としての役目をし、トランザクション・サーバと呼ばれる。作業負荷管理プログラム及びトランザクション・サーバが、同一プロセッサ（コンピュータ）に存在するシステム構成にすることは非常に良いことである。これらのプロセッサはトランザクション・プログラムを実行し、プログラムによって行われるＤＢ要求を処理する。これらの機能は同時実行制御、指標付け、バッファリング、ディスク入出力、その他を有し、及びデータベース層１３乃至１５を有する。本発明の実施例によって仮定されるモデルにおいて、データベースの記録はＢＥＰ全体に区分される。プロセッサB_1で実行しているトランザクションT_1が、ファイル（関連ＤＢ）F_k（１７）のレコードＲ（１８）に関して何らかの処理を要求したものと仮定する。そのレコードがB_1内に存在していれば、その動作は局部的に実行される。さもなければ、動作はレコードを有するプロセッサB_nに機能移管される。著者D. W. Cornell、et al. による題名"On Multisystem Coupling through FunctionRequest Shipping" (IEEE Transactions on Software Engineering、SE-12 (10)、October 1986.) は、上記方法がＩＢＭトランザクション処理用の製品であるＣＩＣＳによってどのように行われるかを述べる。B_nはその動作を処理し、結果とデータをB_1に送信し、B_1はその結果をT_1に送る。ここでの実施例では、ＢＥＰ全体でデータベースを区分していると仮定しているが、経路指定アルゴリズム（及び、本実施例の一部であるスケジューリング・アルゴリズム）は、またデータベースがＢＥＰ間で共有される、共有データベース環境にも適用できる。

7

【００２５】好ましい実施例によれば、データベース管理者は各トランザクション階級において平均応答時間目標を定義する。これらの目標は、g_1、g_2、．．．．、g_xで表される。問題はこれらの目標を達成するためにトランザクション作業負担を管理することである。トランザクション階級の性能目標を定義して実行させる能力は、多様なユーザ・コミュニティによって提示された様々な種類のトランザクションを処理するシステムにとって重要である。さらに処理レベル協定が作業の異なる階級のために必要な応答時間の観点から通常、指定される。著者C. Watson、et al. による題名"The Three Phases of Service Levels" (MainframeJournal、July 1989. ）は、この業界で使用される様々な処理レベル協定の例を与えるので参照されたい。

【００２６】本発明の実施例では２種類の作業負荷管理方法を用いる。第１はＦＥＰで実行されるトランザクション経路指定アルゴリズムの使用である。各々の着信トランザクションはトランザクションによって、要求された推定資源にもとづくＢＥＰの１つに経路指定され、ＢＥＰの現システム状態（ロード）及び応答時間目標への達成度は、その経路選択の結果として満足させられる。第２は動的優先順位方式の使用である。この方式はＢＥＰのスケジューリング優先順位を動的に調整する。これらの優先順位はＣＰＵ処理のために待機中のどのトランザクションをタスク指名するか及び割込みできるか、できないかを決める。

【００２７】経路指定技術について説明する。トランザクション経路指定技術はＦＥＰのデータ通信層で実行される。この経路指定技術は、経路指定を求めるのに必要な３つの集合の制御変数を管理する。これらは次の通りである。

【００２８】１．ベクトルN＝｛N_1、N_2、．．．．、N_x｝：フィールドN_jは現在、システムにある（ＢＥＰへ経路指定された）階級C_jのトランザクション数である＊

8

＊る。

【００２９】２．ベクトルP＝｛P_1、P_2、．．．．、P_x｝：フィールドP_jは階級C_jの性能指標と呼ばれ、階級C_jの現目標達成度を表す。P_jは現平均階級C_jの応答時間を目標g_jによって除算した値である。$P_i \leqq 1$ならば、階級C_jはその目標にかなう。性能指標が小さいほど性能が良好で目標が十分に達成されていることを意味する。ベクトルPを管理する技術は後で説明する。

【００３０】３．Ｓはシステム状態情報を表す。この情報は過去の経路指定決定の経歴、ＢＥＰの現ロード、トランザクション階級の資源使用統計値、その他を有する。この状態情報は経路指定技術の異なる実施によって変わる。性能予測副構成のために必要である状態情報の実施例が後で述べられる。

【００３１】経路指定技術は副機能Ｒ（Ｓ、ｉ、１、ｊ）を用いる。Ｓは現システム状態である。ｉは最後に着信したトランザクションの階級ＩＤ（識別子）である。１はＢＥＰ　ＩＤ、及びｊは任意の階級ＩＤである。この副機能は着信トランザクションT∈C_iがＢＥＰ　B_lに経路指定される場合、既にシステムに存在する階級C_jのトランザクションの平均応答時間を予測する。Ｒの実施例は後で述べる。

【００３２】経路指定技術は各トランザクションがＦＥＰに、着信及び各トランザクションの処理完了毎に呼出される。

【００３３】着信アルゴリズムについて説明する。着信トランザクションT　　C_iがあると仮定する。以下のステップが実行される。

【００３４】１．フィールドT．ａｒｒｉｖａｌ＿ｔｉｍｅの着信時間Tを記録する。

【００３５】２．全ての階級C_j、及び全てのＢＥＰ　B_lを計算する。

【数２】

$$P(j,l) = \alpha \bullet P_j + (1-\alpha) \bullet N_j \bullet \left(\frac{R(S,i,l,j)}{g_j} \right).$$

【００３６】この計算は図５に示されている。値Ｐ（ｊ、１）は、ＴがＢＥＰ　B_lに経路指定された場合における、階級C_jの性能指標の予測された新しい値である。ＴがB_lに経路指定される場合、現在B_lに対してアクティブ（経路指定されている）である階級C_jのトランザクションの応答時間に影響を及ぼす。Ｔは，また他のＢＥＰに対して機能移管を生じさせ、他のＢＥＰに経路指定されている階級C_jのトランザクションの応答時間に影響する。パラメータαは$0 \leqq \alpha \leqq 1$の範囲であり、現システム状態と以前に観測された応答時間（過去の経歴）との比較の重要性に重みをおくのに利用される。

【００３７】３．トランザクションＴは次式を満足するＢＥＰ　B_mに経路指定される。

【数３】

$$\min_{1} \max_{j} \{P(j,1)\}$$

（１を固定してＰ（ｊ，１）をすべてのｊについて計算し、そのときの最大値をＱ（１）とすると、Ｑ（１）を最小にするものをいう。）

【００３８】この手法は全ての性能指標を等しくさせ、且つ可能な限り小さくさせる。この手法によってある階級が目標を大きく上回ったために、他の階級が目標を達成できない、という事態の到来を防止する。

【００３９】４．ＴがB_mに送信された後、経路指定を反映するためにN_j及びＳは更新させられる。

9

【００４０】発信アルゴリズムについて説明する。トランザクションT　　C_i の処理完了（発信）があると仮定する。図６は発信アルゴリズムの図式である。以下のステップが実行される。

【００４１】１．R＝Ｃｕｒｒｅｎｔ＿ｔｉｍｅ－T.ａｒｒｉｖａｌ＿ｔｉｍｅを設定する。

【００４２】２．次式によりP_i の値を更新する

【数４】

$$P_i = \beta \cdot P_i + (1-\beta) \cdot \left(\frac{R}{g_i} \right).$$

【００４３】パラメータβは過去の応答時間と最後に発信した応答時間との相対的な重み付けを示す。その他の技術では最後に発信した階級での平均応答時間を計ることにより、P_i を計算することが可能である。

【００４４】３．N_i 及びSを更新する。

【００４５】動的優先順位付け技術について説明する。ＦＥＰ／ＤＣ層装置は正確に目標達成度をモニタできる、トランザクション処理システムの唯一の構成要素である。従って、ＦＥＰ／ＤＣ層装置はＢＥＰで使用されるトランザクション階級の優先順位を設定できる。ＢＥＰは局所的な情報にもとづく優先順位を設定することはできない。これは、各々のＢＥＰはトランザクション階級内の一部分のみの処理完了をモニタするにすぎないからである。従って、目標達成度に関するＢＥＰ情報は非常に不正確である。

【００４６】動的優先順位付け技術は３つの構成機能から成る。第１の機能は目標達成度にもとづく階級優先順位を更新し、その更新をＢＥＰに知らせるＦＥＰ／ＤＣ機能である。第２の機能はＦＥＰ／ＤＣによって送信された優先順位の更新を受信して処理するＢＥＰの副機能である。第３の機能はＢＥＰ内でスケジューリング手法を実行する副機能である。この第３の副機能はどのトランザクションがＣＰＵを占有するかを決める。これらの３つの機能について次に述べる。

【００４７】ＦＥＰ／ＤＣ更新アルゴリズムについて説明する。ＦＥＰ／ＤＣ更新アルゴリズムはトランザクションが処理完了する度に呼出される。これは前述の発信アルゴリズムのステップ２に続く。このアルゴリズムは内部制御変数ｏ（ｉ）、ｉ＝１、２、．．．．、Ｋを管理する。集合ｏ（ｉ）は集合｛１、２、．．．．、Ｋ｝の置換であり、最大性能指標から最小性能指標にソートされたトランザクション階級ＩＤを表す。言い換えると、集合ｏ（ｉ）は次式の特性を有する。

【数５】$P_{o(i)} \geqq P_{o(i+1)}$

【００４８】処理完了のトランザクションT∈C_i があると仮定する。優先順位更新アルゴリズムは性能指標P_i の更新後、以下のステップを実行する。

【００４９】１．$P_{o(i)} \leqq P_{o(i+1)}$ または$P_{o(i)} \geqq P_{o(i-1)}$ならば、

10

ａ．性能指標の現在値にもとづいて順序、ｏ（１）、ｏ（２）、．．．．、ｏ（Ｋ）を再計算する。

ｂ．ベクトルＰ＝｛P_1、P_2、．．．．、P_k ｝の新しい値を全てのＢＥＰに送信する。

【００５０】順序リストｏ（ｉ）はオーバヘッドを、より低くするのに使用される。Ｂ－ｔｒｅｅなど、性能指標が高い階級を探索するオーバヘッドをより低くする他のいくつかの実施方法がある。階級の相対的順位を実際に変える性能指標の変化だけがＢＥＰに送信される。順位付けを変えない変化は、ＢＥＰの優先順位に影響しないので送信されない。さらにオーバヘッドをより低くするには、ＦＥＰ／ＤＣによって経路指定されたトランザクションの更新された性能指標をピギーバッグ（piggyback＝トランザクションのメッセージの特別なフィールドに挿入）することで可能である。これによってトランザクションがＢＥＰに送られたとき、それとともに新しい性能指標も送ることができるので余分なメッセージが不要となる。本明細書の実施例では、ＦＥＰ／ＤＣがトランザクションをＢＥＰに経路指定する度に、ＢＥＰは最新のベクトルＰが送信されたかをテストする。テスト結果が否定ならば、新しいベクトルＰが経路指定されたトランザクションに付加され、ＢＥＰがトランザクションを受信するときにＢＥＰによって抽出される。この最適化が優先順位を更新するための余分なメッセージの送信の必要性を排除する。オーバヘッドは更新された性能指標ベクトルが送信される間において、処理完了の数を固定することを要求することによってさらに減らすことができる。

【００５１】性能指標ベクトルをＢＥＰに送信する他の代替方法は、新しく経路指定されたトランザクションの階級の優先順位のみ更新することである。例えば、階級C_1のトランザクションT_1が着信すると仮定する。ＦＥＰはT_1をＢＥＰ　B_3 に経路指定し、ＢＥＰ　B_3の階級C_1の優先順位をP_1からP'_1 に変更することを決めたと仮定する。次にメッセージ＜P'_1、T_1＞はＢＥＰ　B_3に送られる。

【００５２】受信更新アルゴリズムについて説明する。各々のＢＥＰはベクトルＰの局所のコピーを管理する。ベクトルＰの新しい値を含むメッセージ（またはトランザクション）を受信後、局所コピーが更新される。

【００５３】ＢＥＰスケジューリング・アルゴリズムについて説明する。ＢＥＰスケジューリング・アルゴリズムは実行中のトランザクションTがＣＰＵを解放して待ち状態に入る度に呼出される。トランザクションはＩ／Ｏ待ち、機能移管応答、ユーザ入力、ロック、その他を含む多くの理由で待ち状態に入る。ＢＥＰスケジューラが割込み可能で、より高位の優先順位である階級に属する他のトランザクションが実行可能になった場合、実行中のトランザクションは待ち状態におかれる。待ち状態が完了したのちに、Ｔは実行準備をさせられ、ＣＰＵに

11

対して待機する。ＢＥＰは全てのトランザクションのトランザクション階級を知っている。機能移管要求の階級は、その要求が実行されるトランザクションの階級と同じである。

【００５４】ＢＥＰスケジューリング・アルゴリズムは少くとも１つの待機トランザクションを有している階級中で最も高位の優先順位を有する階級C_iからトランザクションをタスク指名する。これはラウンド・ロビン・スキャンを使用してトランザクション階級を調べることによって実行される。ラウンド・ロビン・スキャンは同一優先順位を有する他の階級と、階級C_iとを区別するために使用される。階級の優先順位は性能指標によって決まる。一般に階級の性能指標が高くなると、その優先順位が高くなる。しかしながら、性能指標が非常に接近している階級に、厳密な優先順位を与えることを避けるために一定範囲の性能指標を有する階級の一群は同一の優先順位であると規定される。

【００５５】実行中のトランザクションが待ち状態に入る度に、図４にも示されているように以下のステップが実行される。

【００５６】ａ．ｃｏｕｎｔ＝０

【００５７】ｂ．ｉ＝ｌａｓｔ＿ｄｉｓｐａｔｃｈｅｄ

｛ｌａｓｔ＿ｄｉｓｐａｔｃｈｅｄは、スケジューリング・アルゴリズムによってタスク指名された最後のトランザクションの階級ＩＤを記録する。初期設定は１である。｝

【００５８】ｃ．ｂｅｓｔ＿ｃｌａｓｓ＝－１

｛アルゴリズムは、待機トランザクションを有する最高位の優先順位の階級を探すために全てのトランザクション階級を調べる。｝

【００５９】ｄ．ｃｏｕｎｔ≦Ｋの間

１）ｉ＝ｉ＋１を実行する。｛階級はラウンド・ロビンによってスキャンされる。指標が一回りすると、指標は１に設定される。｝

２）もし　ｉ＞Ｋ　ならば

ａ）ｉ＝１

｛アルゴリズムが実行可能状態にあるトランザクション、または調べられた階級がより高位の優先順位を有している場合は、これまでに調べた最高位の優先順位の階級を現階級に設定する。｝

３）ＣＰＵ処理のために待機している階級C_iのトランザクションがある場合、

ａ）もし　ｂｅｓｔ＿ｃｌａｓｓ＝－１　ＯＲ　$P_{best\_class}<P_i-\varepsilon$　ならば｛εは群パラメータである。ε＝０ならば、優先順位は性能指標によって厳密に求められる。さもなければ、優先順位は性能指標がεの範囲内のトランザクションについては一定である。｝

ｉ．ｂｅｓｔ＿ｃｌａｓｓ＝ｉ

４）ｃｏｕｎｔ＝ｃｏｕｎｔ　＋１

｛もうひとつの階級が調べられる｝

12

ＥｎｄＷｈｉｌｅ；｛この時点で、タスク指名される次のトランザクションを有する階級が特定されている。この手法は、単にトランザクションを階級から拾い上げてタスク指名する。｝

【００６０】ｅ．もし　ｂｅｓｔ＿ｃｌａｓｓ≠－１　ならば以下を実行する。

１）待機トランザクションＴを階級C_{best_class}から拾い上げる

２）タスク指名（Ｔ）

３）ｌａｓｔ＿ｄｉｓｐａｔｃｈｅｄ＝ｂｅｓｔ＿ｃｌａｓｓ

【００６１】このアルゴリズムは非割込みのスケジューリング手法を実行する。ＢＥＰスケジューリング・アルゴリズムに対する１つの可能な強化は、強制排除を実行することである。トランザクションＴ∈C_iが実行準備されたと仮定する。全ての実行中のトランザクションは最下位の優先順位のトランザクション　Ｔ‘∈　C_jを見つけるために調べられる。$P_i-\varepsilon>P_j$ならばＴ‘は待機に戻され、Ｔがタスク指名される。

【００６２】トランザクション処理システムのある実施形態において、ＢＥＰは内部マルチプログラミング・レベル制約を実施する。この実施形態において、最大でM_{ik}個のトランザクションがどの時点でもＢＥＰ　B_kでアクティブである。トランザクションはトランザクションがプロセスに連結され、ＣＰＵに対しての待機、ＤＢ要求等が許される限りアクティブであると定義する。非アクティブ・トランザクションはプロセスに連結されアクティブになる前に、アクティブなトランザクションが処理完了するのを単に待つ。マルチプログラミング制約は、通常、メモリ制限及びロック・コンテンションを避けるために考慮される。上記のスケジューリング・アルゴリズムは、またアクティブなトランザクションが処理完了した時に、待機している非アクティブなトランザクションのうちのどれが、アクティブにされるかを決めるために使用される。階級ｉのトランザクションまたは機能移管要求がB_kに着信するときに、B_kにすでにM_{ik}のアクティブなトランザクションがあれば待機させられる。

【００６３】図２は任意のトランザクション・サーバＢＥＰ　B_n内の待ち行列の構造を示す。トランザクション・サーバは複数のＣＰＵ（２つのＣＰＵ２０、２１が示されている）を有することができる。各トランザクション階級は２つの待ち行列を有する。第１の待ち行列はＣＰＵ処理のために待機しているアクティブなトランザクションを含む。第２の待ち行列は処理完了を待つ非アクティブなトランザクションを含む。例えば、階級C_1はＣＰＵ処理を待つ階級C_1の全てのアクティブなトランザクションを含む待ち行列２２と、処理完了を待つ階級C_1からの非アクティブなトランザクションを有する待ち行列２５とを有する。

13

【００６４】応答時間予測装置について説明する。応答時間予測装置は着信が生ずる毎に、経路指定技術によって呼出される副構成要素である。この構成要素はシステムに現在ある全トランザクションが一定の経路指定をされたときにその応答時間の変化を予測する。本明細では、この構成要素の１実施例がＢＥＰ間にて区分されたデータベースを仮定して提案される。共用データ・ベースの場合でも、応答時間予測装置を有することが可能である。他の多数の実施例が可能であり、関連方法が著者C. T. Hsieh, et al. による題名"A Noniterative Approximate SolutionMethod for closed Multichain Queueing Networks"、(PerformanceEvaluation, 9：119-133, 1989.）及び著者P. S. Yu, et al. による題名"Dynamic Transaction Routing in Distributed Database Systems"（IEEETransactions on Software Engineering, SE-14 (9)、September 1988.）に記載されているので参照されたい。本実施例は従来技術に対して、２つの主な新考案がある。第１に、従来技術はＢＥＰにおけるトランザクション階級の優先順位を動的に変更処理しない。第２に、本実施例はほとんどの従来技術よりも低いオーバヘッドを提供する。

【００６５】以下は予測装置への入力である。

ａ．Ｓは現システム状態を示す。 以下は本実施例によって用いられる状態情報を表す。

ｂ．ｉはトランザクション階級ＩＤである。

ｃ．ｌはＢＥＰのＩＤである。

ｄ．ｊはトランザクション階級のＩＤである。

予測装置は階級C_iからの着信トランザクションがＢＥＰ　B_lに経路指定される場合、システムに現在ある階級C_jのトランザクションの平均応答時間を予測し、出力する。

【００６６】状態情報について説明する。本実施例の予測装置によって用いられる状態情報Ｓは、システムに既に存在するトランザクションについての情報を有し、階級及びＢＥＰごとのトランザクション資源消費を予測する。Ｓの第１の構成要素は、システムに既にあるトランザクション数を表す。これは、Ｍ（ｉ、ｌ）で表され、システム中に存在するＢＥＰ　B_lに経路指定されたトランザクションのうち階級C_iのものの数を記録する。この行列は着信トランザクションがＢＥＰに経路指定されるか、またはシステムのトランザクションが処理完了する毎にＦＥＰ／ＤＣによって更新される。

【００６７】他の状態情報はトランザクション階級の資源消費についての統計情報である。この情報はＢＥＰのルーチンをモニタすることによって集められる。各々のＢＥＰは最後に観測された統計値を、複数のＢＥＰによって報告された統計値をマージするＦＥＰ／ＤＣへ定期的に転送できる。通常、トランザクション処理システムのＤＢプロセッサは、トランザクション資源消費に関する統計値を集める。この情報はオフ・ライン性能調整、容量計画、及び原価計算に用いられる。

14

【００６８】ＦＥＰ／ＤＣはＳの一部として以下の資源消費統計値を管理する。

【００６９】ａ．Ｗ（ｉ、ｌ、ｋ）：これはＢＥＰ　B_lに経路指定された階級C_iのトランザクションによってＢＥＰ　B_kで行われたＣＰＵ作業の平均値である。B_lで実行されたＣＰＵ時間はＤＢ要求の処理時間だけでなく、適用業務処理のＣＰＵ時間をも含む。ＣＰＵ作業は、また機能移管及び遠隔のＤＢ要求の処理による$k \neq l$の場合の他のＢＥＰ　B_kでも発生する。この行列の要素は、階級C_iのトランザクションに起因する全てのＣＰＵ作業を含む。

【００７０】ｂ．Ｖ（ｉ、ｌ、ｋ）：この要素は階級C_iのトランザクションがＢＥＰ　B_lに経路指定された、すなわち、ＢＥＰ　B_kを"訪問"した平均回数を記録する。$k \neq l$ならば、その訪問はB_kに対する機能移管のＤＢ要求（準備／実行処理を含む）である。$l = k$の場合、訪問はトランザクションによって生成されたＤＢ要求の数に限定される。この合計は局所及び遠隔での実行と、局所及び遠隔のＤＢの読出し及び書込みを含む。

【００７１】ｃ．Ｉ（ｉ）：これは階級C_iのトランザクションの予測全Ｉ／Ｏ遅延時間である。これはＩ／Ｏサブシステムの遅延時間だけを含む。データが区分化されているために、Ｉ（ｉ）はトランザクションに対する経路指定の決定に依存しない。

【００７２】ｄ．Ｄ（ｉ、ｌ）：これはＢＥＰ　B_lに経路指定された階級C_iのトランザクションの予測全通信遅延時間である。この遅延はＢＥＰを接続しているネットワークを介しての機能移管ＤＢ要求、及び実行／準備メッセージをも含む。

【００７３】ｅ．Ｃ（ｉ、ｌ）：これはB_lに経路指定された階級C_iのトランザクションに対する全ての位置におけるログ（実行／準備）記録の書込みの予測全Ｉ／Ｏ遅延時間である。この遅延時間は標準の２段階の実行プロトコルにおける、様々な最適化のための経路指定の決定に依存する。

【００７４】推定アルゴリズムについて説明する。推定アルゴリズムは予測をするために状態情報を用いる。アルゴリズムはＳの資源消費統計値が更新される毎に、以下の内部行列を予め計算する。

【００７５】ａ．Ｔ（ｉ、ｌ）：これはB_lに経路指定された階級C_iのトランザクションによって消費された資源の和である。これは次式によって与えられる。

【数６】

$$T(i,l)=\left(\sum_{m=1}^{N}W(i,l,m)\right)+D(i,l)+I(i)+C(i,l).$$

【0076】これは待ち合わせ遅延時間がない（すなわち、システムは空である）場合の、B_lに経路指定された階級C_iのトランザクションの応答時間に近似する。
b．B（i、l、k）：これはB_lに経路指定された階級C_iのトランザクションに係わるB_kでの平均ＣＰＵ"バースト長さ"である。これは、次式によって近似する。

【数7】

$$B(i,l,k)=\frac{W(i,l,k)}{V(i,l,k)}.$$

【0077】これらの２つの行列は半静的でシステムの異なる階級のトランザクション数に依存しない。

【0078】応答時間予測装置は図３で概略的に示されているが、これについて説明する。階級ＩＤは$P_1 \geqq P_2 \geqq \cdots \geqq P_K$の順序であると仮定する。これによって説明は簡単になるが、このことはアルゴリズムにとっての本質ではない。アルゴリズムは次に述べる通りである。

【0079】a．内部変数を初期化する。
1）Q（m）＝0；m＝1、2、．．．、Ｎ
{Q（m）はトランザクションがB_mを訪問する場合、ＣＰＵを得るまでトランザクションが待機する時間の量を表す。この量は正確に計算するのは難しく厄介であるので概算値を用いる。アルゴリズムは最高位の優先順位の階級から始まり、Q（m）をゼロに初期化する。アルゴリズムは階級C_iのトランザクションが階級C_l（j＝1、．．．、i）のシステムに現存する全てのトランザクションが処理されるまで待機しなければならないという事実を考慮して、Q（m）を連続して更新する。}
2）ＲＴ（j、k）＝0；j＝1、2、．．．、Ｋ；k＝1、2、．．．、Ｎ
{ＲＴ（j、k）はプロセッサB_kに経路指定された階級jのトランザクションの平均応答時間である。}
3）ＲＰ（j、k、m）＝0；j＝1、2、．．．、Ｋ；k＝1、2、．．．、Ｎ；m＝1、2、．．．、Ｎ
{ＲＰ（j、k、m）は、B_kに経路指定された階級jのトランザクションがＢＥＰ　mに存在する可能性の推定値である。}

a)　$Q(m)=$

$$Q(m)+\left(\sum_{k=1}^{N}M(j,k)\bullet RP(j,k,m)\bullet B(j,k,m)\right).$$

【0080】b．提示された経路指定を反映するために、Ｍ（i、l）の局所のコピーを一時的に増加させる。{様々なトランザクションに対する予測待ち合わせ遅延時間をここで計算する。優先順位の関係から階級C_iのトランザクションは、階級１、２、．．．、j－１からのトランザクションの後にのみ待機する。トランザクションの応答時間は、その処理要求と経験した待ち合わせ遅延時間とで求められる。}

【0081】c．ＦＯＲ　j＝1　ＴＯ　Ｋ　ＤＯ
1）ＦＯＲ　k＝1　ＴＯ　Ｎ　ＤＯ
{ＢＥＰへの訪問に起因する階級／経路指定ごとの待ち合わせ遅延時間分ごとにシステムの期待時間を増加する。}

【数8】

a)　$S(j,k)=$

$$T(j,k)+\sum_{m=1}^{N}V(j,k,m)\bullet Q(m).$$

【0082】{階級C_jのトランザクションがＢＥＰB_mでＣＰＵ処理を待っている予測された確率を更新する。これはB_mにおける処理要求と待ち合わせ遅延時間に比例し、全システム時間に反比例すると考えられる。}
2）ＦＯＲ　k＝1　ＴＯ　Ｎ　ＤＯ
a）ＦＯＲ　m＝1　ＴＯ　Ｎ　ＤＯ

【数9】

i.　$RP(j,k,m)=$

$$\frac{W(j,k,m)+V(j,k,m)\bullet Q(m)}{S(j,k)}$$

【0083】{この階級によって生じた追加の待ち合わせ時間の量だけ、予測待ち合わせ遅延時間を増加させる。これは平均処理バースト長さに対してＣＰＵ処理のために待ち合わせする確率を乗算して近似を得る。}
3）ＦＯＲ　m＝1　ＴＯ　Ｎ　ＤＯ

【数10】

17

【００８４】４）処理要求及び待ち合わせにもとづく、B_kに経路指定されたC_iのトランザクションの応答時間を計算する。

ＦＯＲ　ｋ＝１　ＴＯ　Ｎ　ＤＯ

【数１１】

a)　$$RT(j,k) = T(j,k) + \left(\sum_{m=1}^{N} Q(m) \cdot V(j,k,m) \right).$$

【００８５】ｄ．この時点で、現在システムにある全トランザクションの応答時間が予測された。予測装置によって復帰された平均値は、加重平均として与えられる。

ＦＯＲ　ｊ＝１　ＴＯ　Ｎ　ＤＯ

【数１２】

$$R(S,i,l,j) = \frac{\sum_{k=1}^{N} RT(j,k) \cdot M(j,k)}{N_j}.$$

【００８６】応答時間予測装置の計算コストをより低くできる複数の潜在的な最適化がある。第１はB_lへのC_iの任意の経路指定において、アルゴリズムによって全てのｊに対して、１回の通過でＲ（Ｓ、ｉ、ｌ、ｊ）を計算することが可能である。第２は階級C_iに対しての経路指定の決定は、$P_j > P_i$での何れの階級C_jにおける結果に対して影響を及ぼさない。この基本的な方式に対する多数の他の最適化が可能である。上記で提案されたのとは２つの点において異なる、計算の少ないオーバヘッドであるアルゴリズムを次に述べる。第１はアルゴリズムはＢＥＢＢ$_k$でＣＰＵを待っている階級C_iのトランザクションの確率を計算しない。その代わりに、アルゴリズムは階級C_iのＱ（ｊ、ｋ）によるＢＥＢ　B_kの作業量を計算し、トランザクションがＣＰＵに連結される前のトランザクションがＢＥＰＢ$_k$で待たなければならない時間量が、Ｑ（１、ｋ）＋．．．＋Ｑ（ｉ、ｋ）であると仮定する。第２はアルゴリズムは全ての階級の経路指定の効果を予測しない。その代わりに、トランザクションの推定応答時間が最小であるＢＥＰに対してトランザクションを経路指定する。

【００８７】簡素化アルゴリズムについて説明する。前述のように、階級C_iのトランザクションはＦＥＰに着信し、トランザクションがB_iに経路指定された場合にトランザクションの応答時間に対する推定が所望されると仮定した。ここでもトランザクション指標はｊ＞ｉの場合にトランザクションｊがトランザクションｉより高位の優先順位を持つように、順序付けされると仮定する。

18

【００８８】推定アルゴリズムについて説明する。

ａ．内部変数を初期化する。

１）Ｑ（ｊ、ｋ）＝０：　ｊ＝１、．．．、Ｋ、　ｋ＝１、．．．、Ｎ

ｂ．　ＦＯＲ　ｊ＝１　ＴＯ　ｉ　ＤＯ

１）ＦＯＲ　ｋ＝１　ＴＯ　Ｎ　ＤＯ｛階級C_iによるB_kの作業量を予測する。｝

【数１３】

$$Q(j,k) = \sum_{m=1}^{N} M(j,m) \cdot W(j,m,k)$$

【００８９】ｃ．｛着信トランザクションがB_iに経路指定される場合、着信トランザクションの応答時間を計算する。｝

【数１４】

$$RT(i,l) = T(i,l) + \sum_{k=1}^{N} V(i,l,k) \left(\sum_{j=1}^{i} Q(j,k) \right)$$

【００９０】経路指定アルゴリズムを説明する。

ａ．トランザクションは、ＲＴ（ｉ、ｋ）が最小となるように、B_iに経路指定される。

ｂ．Ｍ（ｉ、１）＝Ｍ（ｉ、１）＋１｛B_iにおけるC_iのトランザクション数を１つ増加させる。｝

【００９１】本発明は、何れの多重プロセッサ・トランザクション処理システムで実行できる。本明細に記述された本発明は、区分されたデータ（共有なし）のデータベース・システムに最も直接に適用できる。本発明は応答時間予測装置を修正することによって、データ共有アーキテクチャにも適用できる。トランザクション経路指定技術及び動的優先順位付け技術は、データ共有アーキテクチャに対して変わることはない。本発明は又、周期的に性能情報を共有するＤＣを有することによって、複数のＦＥＰ／ＤＣのトランザクション処理システムにも応用できる。

【００９２】本発明は、好ましい実施例に関して特に図示して説明を行ったが、当業者は本発明の趣旨、及び範囲内で形式、及び詳細を様々に変更できることが理解できよう。

【００９３】

【発明の効果】ＢＥＰにおけるトランザクション階級の優先順位を動的に変更処理し、ほとんどの従来技術よりも低いオーバヘッドを提供する。

【図面の簡単な説明】

【図１】本発明の好ましい実施例であるシステムのブロック図である。

【図２】トランザクション・サーバにおいて、マルチプログラミング・レベルに入るために、またはＣＰＵ処理

をするために待機するトランザクションの様々な待ち行列を示すブロック図である。

【図３】好ましい実施例に従って現平均階級応答時間の計算を例示する図である。

【図４】好ましい実施例のトランザクション・サーバによって使用されるスケジューリング・アルゴリズムの流れ図である。

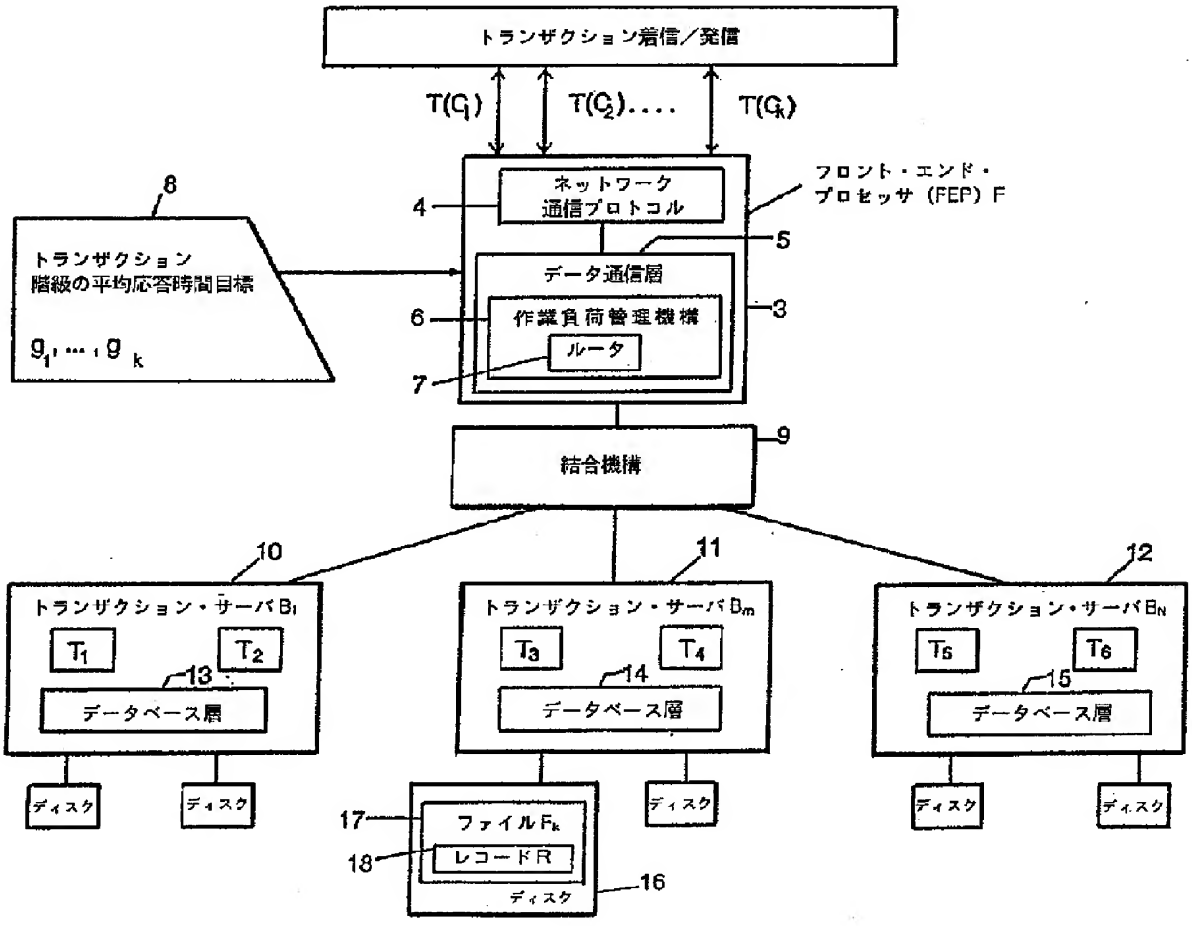
【図５】着信トランザクションに対する新しい階級性能指標の好ましい実施例による予測を例示する図である。

【図６】トランザクションの処理完了時における、現平均階級応答時間の好ましい実施例の更新を例示する図である。

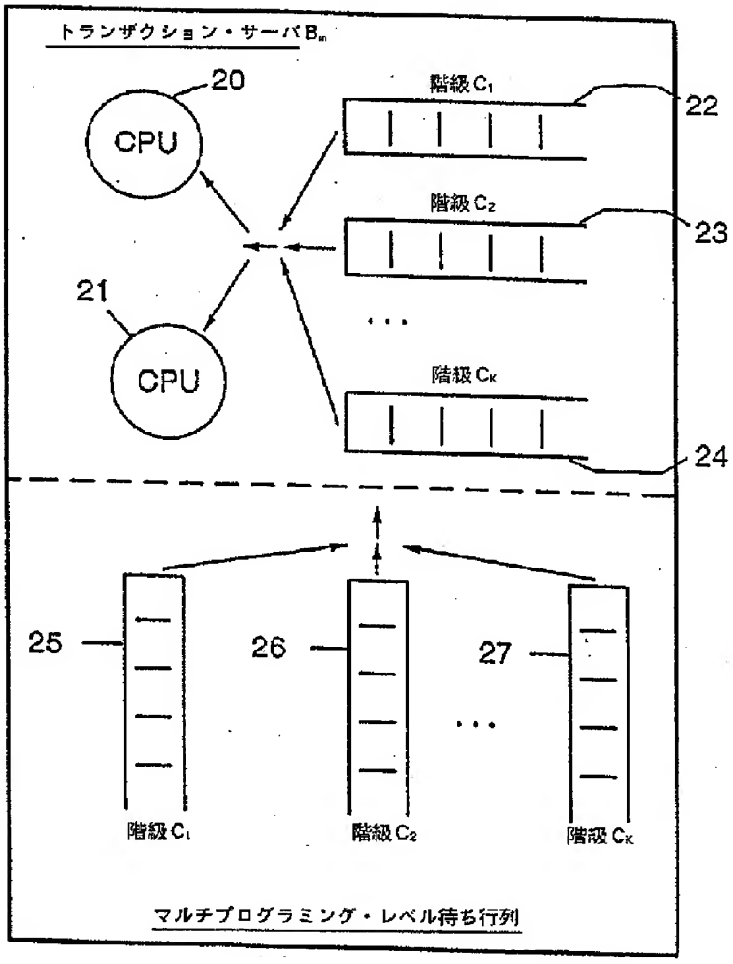
【符号の説明】

３　フロント・エンド・プロセッサ（ＦＥＰ）Ｆ
４　ネットワーク通信プロトコル
５　データ通信層
６　作業負荷管理プログラム
７　ルータ
９　結合機構
１０　トランザクション・サーバB_1
１１　トランザクション・サーバB_m
１２　トランザクション・サーバB_N
１７　ファイルF_k
１８　レコードＲ
２２　階級C_1

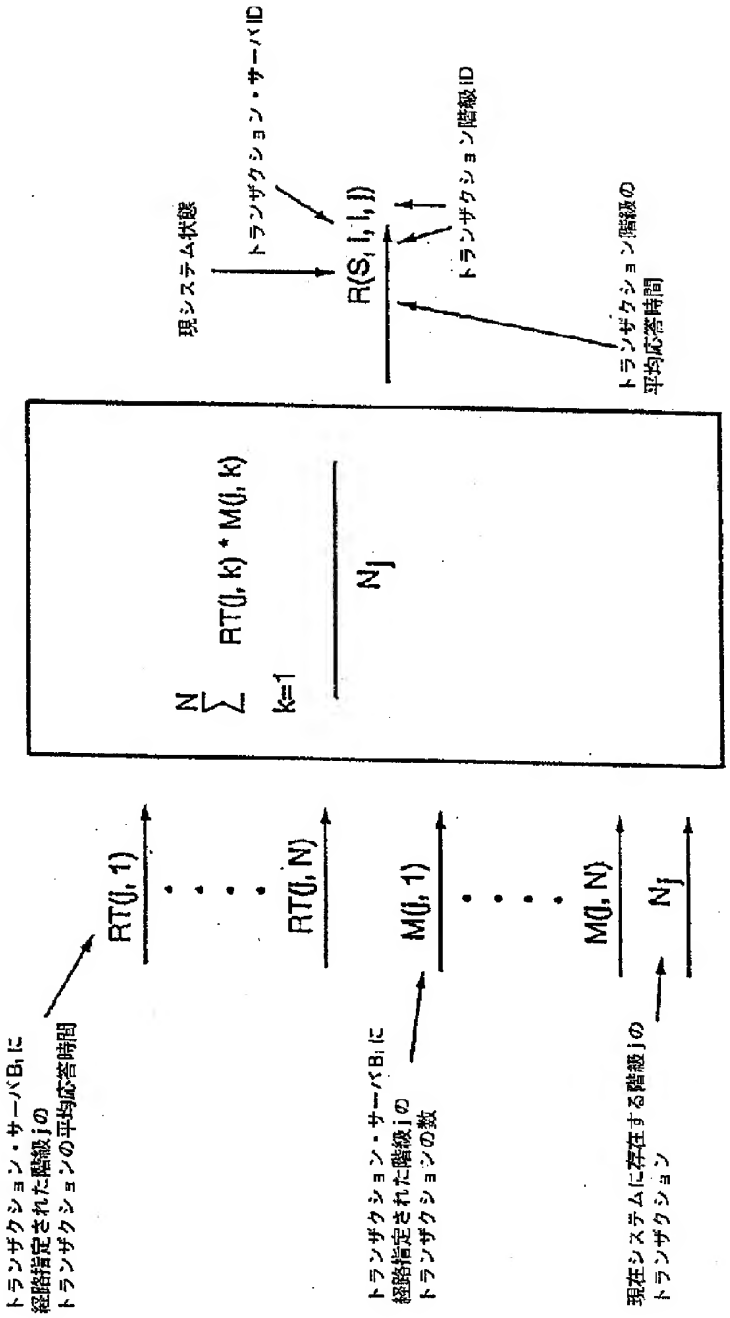
【図１】



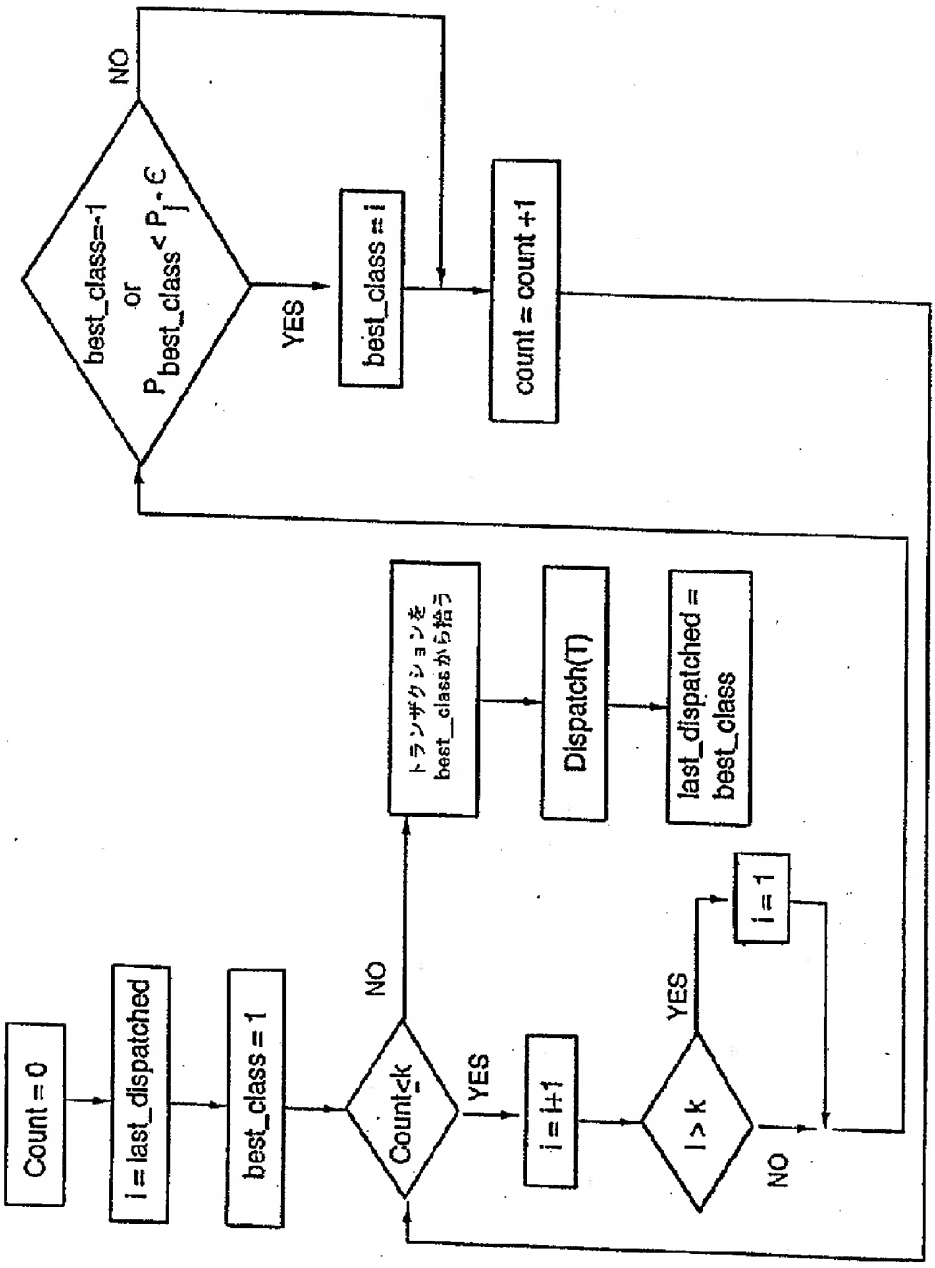
【図２】


トランザクション・サーバB_m
20
CPU
21
CPU
階級C_1
22
階級C_2
23
階級C_K
24
25
26
27
階級C_1
階級C_2
階級C_K
マルチプログラミング・レベル待ち行列

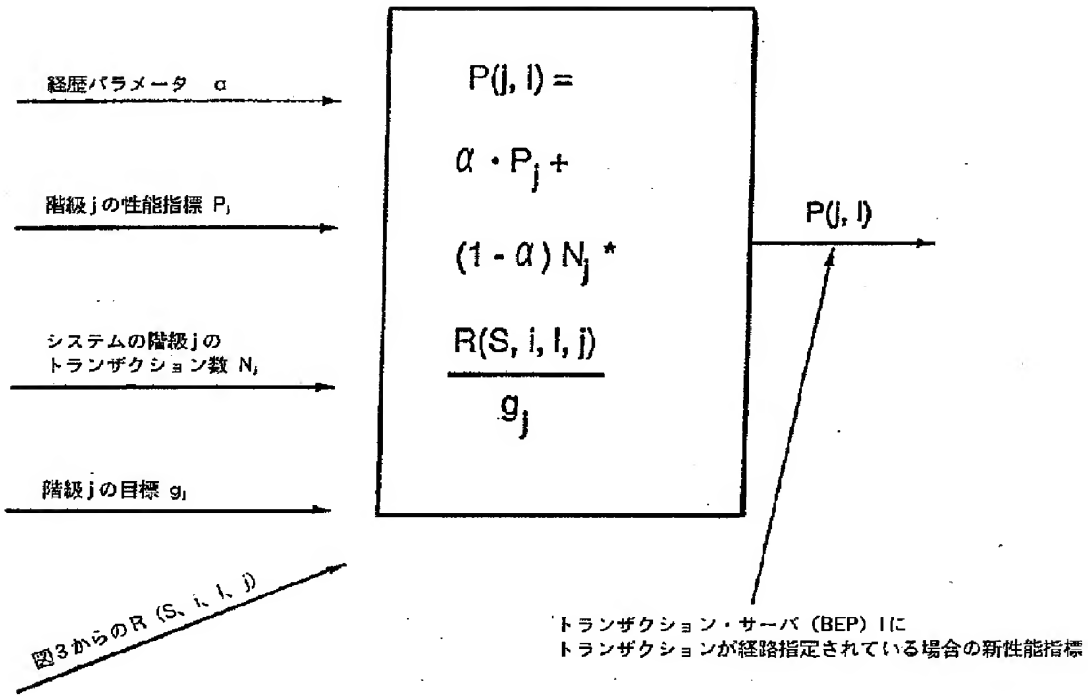
【図3】


現システム状態
トランザクション・サーバID
R(S, I, J)
トランザクション階級ID
トランザクション階級の
平均応答時間
N
Σ RT(I, k) * M(I, k)
k=1
Nj
トランザクション・サーバBIに
経路指定された階級Jの
トランザクションの平均応答時間
RT(I, 1)
RT(I, N)
トランザクション・サーバBIに
経路指定された階級Jの
トランザクションの数
M(I, 1)
M(I, N)
現在システムに存在する階級Jの
トランザクション
Nj

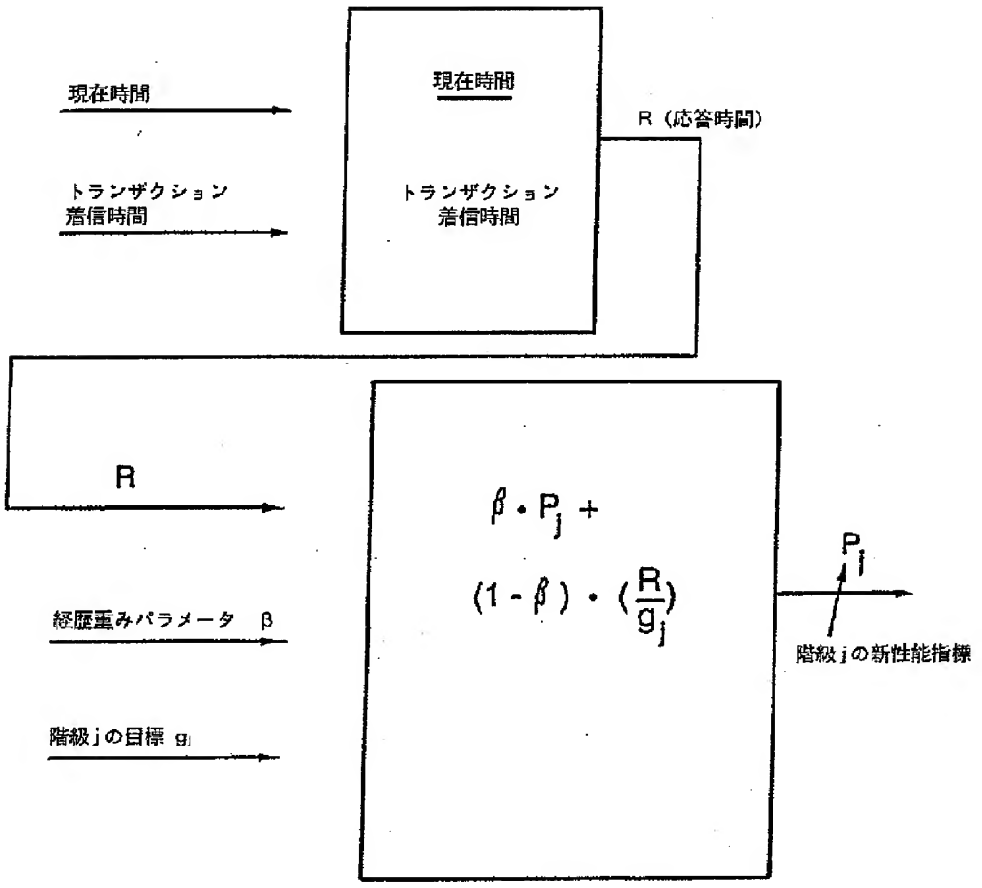
【図４】


Count = 0
i = last_dispatched
best_class = 1
Count<k
NO
YES
i = i+1
i > k
YES
NO
i = 1
トランザクションを best_class から始コ
Dispatch(T)
last_dispatched = best_class
best_class=-1 or P best_class < P j * C
NO
YES
best_class = i
count = count +1

【図５】


経歴パラメータ　α
階級ｊの性能指標 P_j
システムの階級ｊの
トランザクション数 N_j
階級ｊの目標 g_j
図3からのR (S、i、l、j)
$P(j, l) = \alpha \cdot P_j + (1 - \alpha) N_j * \frac{R(S, i, l, j)}{g_j}$
P(j, l)
トランザクション・サーバ（BEP）Ｉに
トランザクションが経路指定されている場合の新性能指標

【図６】



フロントページの続き

(72)発明者　レオニダス・ジョージアディス
アメリカ合衆国10514、ニューヨーク州チャパクア、オールド・ミル・ロード　38

(72)発明者　クリストス・ニコラス・ニコロー
アメリカ合衆国10023、ニューヨーク州ニューヨーク、アパートメント　12イー、リバーサイド・ドライブ　5